

OIPE

RAW SEQUENCE LISTING

DATE: 10/18/2001

PATENT APPLICATION: US/09/925,796

TIME: 11:40:52

Input Set : N:\Crf3\RULE60\09925796.txt

Output Set: N:\CRF3\10182001\I925796.raw

#2.

3 <110> APPLICANT: Case, Casey C.
 4 Zhang, Lei
 5 Sangamo Biosciences, Inc.
 7 <120> TITLE OF INVENTION: Functional Genomics Using Zinc Finger Proteins
 9 <130> FILE REFERENCE: 019496-002000US
 11 <140> CURRENT APPLICATION NUMBER: 09/925,796
 12 <141> CURRENT FILING DATE: 2001-08-09
 14 <150> PRIOR APPLICATION NUMBER: 09/395,448
 15 <151> PRIOR FILING DATE: 1999-09-14
 17 <150> PRIOR APPLICATION NUMBER: 09/229,037
 18 <151> PRIOR FILING DATE: 1999-01-12
 20 <160> NUMBER OF SEQ ID NOS: 23
 22 <170> SOFTWARE: PatentIn Ver. 2.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 25
 26 <212> TYPE: PRT
 27 <213> ORGANISM: Artificial Sequence
 29 <220> FEATURE:
 30 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif
 31 of C2H2 class of zinc finger proteins (ZFP)
 33 <220> FEATURE:
 34 <221> NAME/KEY: MOD_RES
 35 <222> LOCATION: (2)..(3)
 36 <223> OTHER INFORMATION: Xaa = any amino acid
 38 <220> FEATURE:
 39 <221> NAME/KEY: MOD_RES
 40 <222> LOCATION: (4)..(5)
 41 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
 43 <220> FEATURE:
 44 <221> NAME/KEY: MOD_RES
 45 <222> LOCATION: (7)..(18)
 46 <223> OTHER INFORMATION: Xaa = any amino acid
 48 <220> FEATURE:
 49 <221> NAME/KEY: MOD_RES
 50 <222> LOCATION: (20)..(22)
 51 <223> OTHER INFORMATION: Xaa = any amino acid
 53 <220> FEATURE:
 54 <221> NAME/KEY: MOD_RES
 55 <222> LOCATION: (23)..(24)
 56 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
 58 <400> SEQUENCE: 1
 W--> 59 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 60 1 5 10 15
 W--> 62 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His
 63 20 25
 66 <210> SEQ ID NO: 2
 67 <211> LENGTH: 10

FATTEN

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/925,796

DATE: 10/18/2001

TIME: 11:40:52

Input Set : N:\Crf3\RULE60\09925796.txt

Output Set: N:\CRF3\10182001\I925796.raw

68 <212> TYPE: DNA
 69 <213> ORGANISM: Artificial Sequence
 71 <220> FEATURE:
 72 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
 73 with two overlapping D-able subsites
 75 <220> FEATURE:
 76 <221> NAME/KEY: modified_base
 77 <222> LOCATION: (1)..(2)
 78 <223> OTHER INFORMATION: n = g, a, c or t
 80 <220> FEATURE:
 81 <221> NAME/KEY: modified_base
 82 <222> LOCATION: (5)
 83 <223> OTHER INFORMATION: n = g, a, c or t
 85 <220> FEATURE:
 86 <221> NAME/KEY: modified_base
 87 <222> LOCATION: (8)
 88 <223> OTHER INFORMATION: n = g, a, c or t
 90 <220> FEATURE:
 91 <221> NAME/KEY: modified_base
 92 <222> LOCATION: (9)
 93 <223> OTHER INFORMATION: n = a, c or t; if g, then position 10 cannot be g
 94 or t
 96 <400> SEQUENCE: 2

W--> 97 nngkngknnn

10

100 <210> SEQ ID NO: 3
 101 <211> LENGTH: 10
 102 <212> TYPE: DNA
 103 <213> ORGANISM: Artificial Sequence
 105 <220> FEATURE:
 106 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
 107 with three overlapping D-able subsites
 109 <220> FEATURE:
 110 <221> NAME/KEY: modified_base
 111 <222> LOCATION: (1)..(2)
 112 <223> OTHER INFORMATION: n = g, a, c or t
 114 <220> FEATURE:
 115 <221> NAME/KEY: modified_base
 116 <222> LOCATION: (5)
 117 <223> OTHER INFORMATION: n = g, a, c or t
 119 <220> FEATURE:
 120 <221> NAME/KEY: modified_base
 121 <222> LOCATION: (8)
 122 <223> OTHER INFORMATION: n = g, a, c or t
 124 <400> SEQUENCE: 3

W--> 125 nngkngkngk

10

128 <210> SEQ ID NO: 4
 129 <211> LENGTH: 5
 130 <212> TYPE: PRT
 131 <213> ORGANISM: Artificial Sequence

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/925,796

DATE: 10/18/2001

TIME: 11:40:52

Input Set : N:\Crif3\RULE60\09925796.txt

Output Set: N:\CRF3\10182001\I925796.raw

133 <220> FEATURE:
134 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
136 <400> SEQUENCE: 4
137 Asp Gly Gly Gly Ser
138 1 5
141 <210> SEQ ID NO: 5
142 <211> LENGTH: 5
143 <212> TYPE: PRT
144 <213> ORGANISM: Artificial Sequence
146 <220> FEATURE:
147 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
149 <400> SEQUENCE: 5
150 Thr Gly Glu Lys Pro
151 1 5
154 <210> SEQ ID NO: 6
155 <211> LENGTH: 9
156 <212> TYPE: PRT
157 <213> ORGANISM: Artificial Sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
162 <400> SEQUENCE: 6
163 Leu Arg Gln Lys Asp Gly Glu Arg Pro
164 1 5
167 <210> SEQ ID NO: 7
168 <211> LENGTH: 4
169 <212> TYPE: PRT
170 <213> ORGANISM: Artificial Sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
175 <400> SEQUENCE: 7
176 Gly Gly Arg Arg
177 1
180 <210> SEQ ID NO: 8
181 <211> LENGTH: 5
182 <212> TYPE: PRT
183 <213> ORGANISM: Artificial Sequence
185 <220> FEATURE:
186 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
188 <400> SEQUENCE: 8
189 Gly Gly Gly Gly Ser
190 1 5
193 <210> SEQ ID NO: 9
194 <211> LENGTH: 8
195 <212> TYPE: PRT
196 <213> ORGANISM: Artificial Sequence
198 <220> FEATURE:
199 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
201 <400> SEQUENCE: 9
202 Gly Gly Arg Arg Gly Gly Gly Ser

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/925,796

DATE: 10/18/2001

TIME: 11:40:52

Input Set : N:\Crif3\RULE60\09925796.txt

Output Set: N:\CRF3\10182001\I925796.raw

```

203   1           5
206 <210> SEQ ID NO: 10
207 <211> LENGTH: 9
208 <212> TYPE: PRT
209 <213> ORGANISM: Artificial Sequence
211 <220> FEATURE:
212 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
214 <400> SEQUENCE: 10
215 Leu Arg Gln Arg Asp Gly Glu Arg Pro
216   1           5
219 <210> SEQ ID NO: 11
220 <211> LENGTH: 12
221 <212> TYPE: PRT
222 <213> ORGANISM: Artificial Sequence
224 <220> FEATURE:
225 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
227 <400> SEQUENCE: 11
228 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
229   1           5           10
232 <210> SEQ ID NO: 12
233 <211> LENGTH: 16
234 <212> TYPE: PRT
235 <213> ORGANISM: Artificial Sequence
237 <220> FEATURE:
238 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
240 <400> SEQUENCE: 12
241 Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
242   1           5           10           15
245 <210> SEQ ID NO: 13
246 <211> LENGTH: 97
247 <212> TYPE: PRT
248 <213> ORGANISM: Artificial Sequence
250 <220> FEATURE:
251 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP sequence in
252     control construct
254 <400> SEQUENCE: 13
255 Val Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly
256   1           5           10           15
258 Lys Val Tyr Gly Gly His Asp Thr Val Val Gly His Leu Arg Trp His
259           20           25           30
261 Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg
262           35           40           45
264 Phe Thr Ala Ala Asp Glu Val Gly Leu His Lys Arg Thr His Thr Gly
265           50           55           60
267 Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Leu Val
268   65           70           75           80
270 Val Ala Thr Gln Leu His Ile Lys Thr His Gln Asn Lys Lys Gly Gly
271           85           90           95
273 Ser

```


RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/925,796

DATE: 10/18/2001

TIME: 11:40:52

Input Set : N:\Crif3\RULE60\09925796.txt

Output Set: N:\CRF3\10182001\I925796.raw

```

277 <210> SEQ ID NO: 14
278 <211> LENGTH: 292
279 <212> TYPE: DNA
280 <213> ORGANISM: Artificial Sequence
282 <220> FEATURE:
283 <223> OTHER INFORMATION: Description of Artificial Sequence:designed ZFP
284     construct (from KpnI to BamHI) targeting 9-base
285     pair target site in VEGF promoter
287 <220> FEATURE:
288 <221> NAME/KEY: CDS
289 <222> LOCATION: (2)..(292)
291 <400> SEQUENCE: 14
292 g gta ccg ggc aag aag aag cag cac atc tgc cac atc cag ggc tgt ggt 49
293   Val Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly
294     1           5           10           15
296 aaa gtt tac ggc cgc tcc gac aac ctg acc cgc cac ctg cgc tgg cac 97
297 Lys Val Tyr Gly Arg Ser Asp Asn Leu Thr Arg His Leu Arg Trp His
298           20           25           30
300 acc ggc gag agg cct ttc atg tgt aca tgg tcc tac tgt ggt aaa cgc 145
301 Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg
302           35           40           45
304 ttc acc aac cgc gac acc ctg gcc cgc cac aag cgt acc cac acc ggt 193
305 Phe Thr Asn Arg Asp Thr Leu Ala Arg His Lys Arg Thr His Thr Gly
306           50           55           60
308 gag aag aaa ttt gct tgt ccg gaa tgt ccg aag cgc ttc atg cgc tcc 241
309 Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser
310 65           70           75           80
312 gac cac ctg tcc aag cac atc aag acc cac cag aac aag aag ggt gga 289
313 Asp His Leu Ser Lys His Ile Lys Thr His Gln Asn Lys Lys Gly Gly
314           85           90           95
316 tcc 292
317 Ser
320 <210> SEQ ID NO: 15
321 <211> LENGTH: 97
322 <212> TYPE: PRT
323 <213> ORGANISM: Artificial Sequence
W--> 324 <220> FEATURE:
324 <223> OTHER INFORMATION: Description of Artificial Sequence:designed ZFP
328 <400> SEQUENCE: 15
329 Val Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly
330   1           5           10           15
332 Lys Val Tyr Gly Arg Ser Asp Asn Leu Thr Arg His Leu Arg Trp His
333           20           25           30
335 Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg
336           35           40           45
338 Phe Thr Asn Arg Asp Thr Leu Ala Arg His Lys Arg Thr His Thr Gly
339           50           55           60
341 Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser
342 65           70           75           80

```


VERIFICATION SUMMARY

PATENT APPLICATION: US/09/925,796

DATE: 10/18/2001

TIME: 11:40:53

Input Set : N:\Crf3\RULE60\09925796.txt

Output Set: N:\CRF3\10182001\I925796.raw

L:59 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:62 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:97 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:125 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:324 M:258 W: Mandatory Feature missing, <220> FEATURE: